

# CS544 HW3

---

## Task 1: Vocabulary Creation

---

**What is the selected threshold for unknown words replacement?**

2

**What is the total size of your vocabulary and what is the total occurrences of the special token '< unk >' after replacement?**

Total size of my vocabulary:43195

Total occurrences of the special token <unk>:20011

## Task 2: Model Learning

---

**How many transition and emission parameters in your HMM?**

Total number of transition parameters:1398

Total number of emission parameters:50319

## Task 3: Greedy Decoding with HMM

---

Greedy decoding accuracy on dev data:94.3%

## Task 4: Viterbi Decoding with HMM

---

Viterbi decoding accuracy on dev data:95.4%

**#Explanation about my solution is written in the comment of the code file**